# A Solution Theory for Quasilinear Singular SPDEs

MÁTÉ GERENCSÉR
*IST Austria*

MARTIN HAIRER
*Imperial College London*

**Abstract**

We give a construction allowing us to build local renormalized solutions to general quasilinear stochastic PDEs within the theory of regularity structures, thus greatly generalizing the recent results of [1, 5, 11]. Loosely speaking, our construction covers quasilinear variants of all classes of equations for which the general construction of [3, 4, 7] applies, including in particular one-dimensional systems with KPZ-type nonlinearities driven by space-time white noise. In a less singular and more specific case, we furthermore show that the counterterms introduced by the renormalization procedure are given by local functionals of the solution. The main feature of our construction is that it allows exploitation of a number of existing results developed for the semilinear case, so that the number of additional arguments it requires is relatively small. 

## 1 Introduction

Amidst the recent heightened interest in singular stochastic partial differential equations (SPDEs), three different methods [1,5,11] have been developed to extend the theory to quasilinear equations. The first two of these worked with paracontrolled calculus, while [11] introduced a new variation of previous techniques to treat singular SPDEs that is closer to the theories of rough paths and regularity structures, but flexible enough to cover quasilinear variants. For a comparison between them in terms of scope we refer the reader to the introduction of [5], but it should be noted that in a sense all of them deal with the "first interesting" case, when the noise is just barely too rough for the product $a(u)\Delta u$ to make sense. In particular, quasilinear variants of the KPZ equation, or for example the parabolic Anderson model in a generalized form in three dimensions, are all outside of the scope of these works. One exception is the forthcoming work [10], which extends [11] to the next regime of regularity and which includes noises slightly better than space-time white noise in one dimension (similar to the setting of our example (1.1) below).

Communications on Pure and Applied Mathematics, Vol. LXXII, 1983–2005 (2019)

In the present article, we tackle this problem within the framework of regularity structures. The generality in which we succeed in building local solution theories is, in some sense, optimal: loosely speaking, we show that if an equation can be solved with regularity structures and its solution has positive regularity, then its quasilinear variants can also be solved (locally). We deal with both the analytic and the probabilistic side of the theory in the sense that we show that the general machinery developed in [4] can be exploited in order to produce random models that do precisely fit our needs. Another major advantage of our approach is that its formulation is such that it allows leveraging many existing results from the semilinear situation without requiring us to reinvent the wheel. This is why, despite its much greater generality, this article is significantly shorter than the works mentioned above.

The only disadvantage of our approach, compared to [1, 5, 11], is that it is not obvious at all a priori why the counterterms generated by the renormalization procedure should be local in the solution. The reason for this is that our method relies on the introduction of additional "nonphysical" components to our equation, which are given by some nonlocal, nonlinear functionals of the solution, and we cannot rule out in general that the counterterms depend on these nonphysical terms. We do, however, address the question of the precise form of the counterterms in a relatively simple case. For equation (1.1) we verify that, provided that the renormalization constants are chosen in a specific way, all nonlocal contributions cancel out exactly. This specific choice of constants happens to differ in general from the BPHZ renormalization introduced in [3], but does still allow showing convergence of the underlying renormalized model. The reason for a lack of a general statement is that the algebraic machinery developed in [2, 4], which allows showing that counterterms are always local in the semilinear case, does not appear to be applicable in a simple way. However, we do expect that this is something that could be addressed in future work.

The concrete example we consider is a slightly regularized version of the quasilinear variant of the generalized KPZ equation, which formally reads as

$$(\partial_t - a(u)\partial_x^2)u = F_0(u)(\partial_x u)^2 + F_1(u)\xi, \quad \text{on } (0,1]\times\mathbb{T},\ u(0,\cdot) = u_0, \tag{1.1}$$

where $\xi$ is a translation-invariant Gaussian noise on $\mathbb{R}\times\mathbb{T}$ with covariance function $\mathscr{C}$ satisfying $|\mathscr{C}(t,x)| \le (|t|^{1/2} + |x|)^{-3+\nu}$ for some $\nu > 0$, $u_0 \in \mathcal{C}^{\overline{\nu}}$ for some $\overline{\nu} > 0$, $a$ is a smooth function taking values in $\mathfrak{K}$ for some compact $\mathfrak{K} \subset (0,\infty)$, and $F_0$ and $F_1$ are smooth functions. The quasilinear equations considered in previous works [1, 5, 11] correspond to situations where $\nu > \frac{1}{3}$, $F_0 = 0$. Let us take a compactly supported, nonnegative, symmetric (under the involution $x \mapsto -x$) smooth function $\rho$ integrating to 1, set $\rho^\varepsilon(t,x) = \varepsilon^{-3}\rho(\varepsilon^{-2}t, \varepsilon^{-1}x)$, and define $u^\varepsilon$

as the classical solution of

$$\begin{aligned}(1.2)\qquad &(\partial_t - a(u^\varepsilon)\partial_x^2)u^\varepsilon \\ &\quad = F_0(u^\varepsilon)(\partial_x u^\varepsilon)^2 + F_1(u^\varepsilon)(\rho^\varepsilon * \xi) \\ &\qquad - C^\varepsilon_{a(u^\varepsilon)}(aF_1'F_1 - a'F_1^2 + F_1^2F_0)(u^\varepsilon), \quad u(0,\cdot) = u_0,\end{aligned}$$

where $C^\varepsilon_c$ is some smooth function of $c \in \mathfrak{K}$. We then have the following (renormalized) well-posedness result for equation (1.1), which will be proved in Section 5.

THEOREM 1.1. *There exist deterministic smooth functions $C^\varepsilon_\cdot$ such that for all $u_0 \in \mathcal{C}^{\overline{\nu}}$ there exists a random time $\tau > 0$ such that $u^\varepsilon$ converges in probability in $\mathcal{C}([0,\tau] \times \mathbb{T}) \cap \mathcal{C}^{1/2}_{\text{loc}}((0,\tau] \times \mathbb{T})$ to a limit $u$. Furthermore, with a suitable choice of $C^\varepsilon_\cdot$, one can ensure that the limit $u$ is independent of $\rho$.*

*Remark* 1.2. We would like to stress again that we only need the condition $\nu > 0$ in order to guarantee that the counterterms created by the renormalization of the underlying model are local functions of $u$. The rest of the argument works down to $\nu > -\frac{1}{2}$ (including in particular the case of space-time white noise), at which point the conditions of [4, thm. 2.14] are violated, and one expects a qualitative change of the scaling behavior of the solution. Similarly, we consider a scalar equation driven by a single noise purely for the sake of notational convenience. The exact same proof also applies for example to systems of the type

$$\partial_t u_i = a_{ij}(u)\partial_x^2 u_j + F^{(2)}_{ijk}(u)(\partial_x u_j)(\partial_x u_k) + F^{(1)}_{ij}(u)(\partial_x u_j) + F^{(0)}_{ij}(u)\xi_j,$$

with $a$ taking values in some compact set of strictly positive definite symmetric matrices and implicit summation over $j, k$.

The structure of the remainder of this article goes as follows. In Section 2, we first give an equivalent formulation of a general quasilinear SPDE that is the main remark this article is based on. The main purpose of this reformulation is to write the equation in integral form in a way that resembles the mild formulation for semilinear problems. In particular, this can be done in such a way that the product $a(u) \cdot \partial_x^2 u$ never appears and is replaced instead by seemingly more complicated terms that however exhibit better scaling/regularity properties. In Section 3, we then show how to build a suitable regularity structure that allows the formulation of the fixed point problem derived in Section 2. This is very similar to what is done in [3, 7] with the unusual twist that each symbol represents an *infinite-dimensional* subspace of the resulting regularity structure rather than a one-dimensional one. The formulation of the fixed point problem is then done in Section 4. Finally, we treat a concrete example in Section 5, where we also verify "by hand" that in this case the renormalization procedure does indeed produce only local counterterms.

## 2 An Equivalent Formulation

The main observation on which this article builds is that, at least for smooth drivers/solutions, a quasilinear equation is equivalent to another equation whose principal (smoothing) part does make sense even in the limit when the driving noise is taken to be rough. The right-hand side of this new equation may however exhibit ill-defined products (sometimes even if the original right-hand side did not), but that situation is already closer to the ones that the theory of [7] was developed for.

To describe this alternative formulation, we restrict our attention to the case of perturbations of the heat equation on the one-dimensional torus $\mathbb{T}$, but it is straightforward to generalize this to other situations. In this case, one wants to solve initial value problems of the type

$$\big(\partial_t - a(u)\partial_x^2\big)u = F(u,\xi) \quad \text{on } (0,1]\times\mathbb{T}, \;\; u(0,\cdot) = u_0, \tag{2.1}$$

where $a$ is a smooth function taking values in $\mathfrak{K}$ for some compact $\mathfrak{K} \subset (0,\infty)$, $F$ is a subcritical (in the sense of [3, 7]) local nonlinearity, and $\xi$ is a noise term, which for (2.1) to make sense, is assumed to be smooth for the moment.

*Remark* 2.1. Assuming that we are interested in noises $\xi \in \mathcal{C}^{\alpha-2}$ for $\alpha \in (0,1)$, so that potential solutions are expected to be of class $\mathcal{C}^\alpha$, $F$ being subcritical is equivalent to assuming that it is of the form

$$F(u,\xi) = F_0(u,\partial_x u) + F_1(u)\xi,$$

where $F_0\colon \mathbb{R}^2 \to \mathbb{R}$ and $F_1\colon \mathbb{R}\to\mathbb{R}$ are smooth functions and the dependence of $F_0$ in its second argument is polynomial of degree strictly less than $(2-\alpha)/(1-\alpha)$.

It will be convenient to write the equation in a more global way: setting $f = \mathbf{1}_{t>0}F(u,\xi) + \delta\otimes u_0$, where $\delta$ is the Dirac mass at time 0 and both $F$ and $u_0$ are extended periodically to all of $\mathbb{R}$, (2.1) is equivalent to

$$\big(\partial_t - a(u)\partial_x^2\big)u = f \quad \text{on } (-\infty,1]\times\mathbb{R}. \tag{2.2}$$

For $c>0$, denote by $P(c,\cdot)$ the Green's function of $\partial_t - c\partial_x^2$ on $\mathbb{T}$. Note that $P$ is smooth as a function of $c$ away from the origin and one has the identity

$$\tfrac{\partial^\ell}{\partial c^\ell}P(c,\cdot) = \partial_x^{2\ell}\,\underbrace{P(c,\cdot)*\cdots*P(c,\cdot)}_{\ell+1 \text{ times}}, \tag{2.3}$$

where the convolutions are in space-time. Introduce operators $I_\ell^{(k)}$ acting on smooth functions $b$ and $f$ by

$$I_\ell^{(k)}(b,f)(z) = \int \big(\partial_x^k\partial_c^\ell P\big)(b(z),z-z')f(z')\,dz'. \tag{2.4}$$

We will also use the shorthand expressions $I = I_0^{(0)}$, $I' = I_0^{(1)}$, $I_1 = I_1^0$, etc. Note that these operators are linear in their second argument but not in the first.

Note also that, by a simple integration by parts, one has the identities

$$I_\ell^{(k+m)}(b, f)(z) = I_\ell^{(k)}(b, \partial_x^m f)(z).$$

Although $I(b, f)$ is of course not the same as the solution map to $(\partial_t - b\partial_x^2)u = f$ if $b$ is nonconstant, it turns out that solving (2.2) is equivalent to solving an equation of the type

$$u = I(a(u), \hat{f}), \tag{2.5a}$$

where $\hat{f} = \mathbf{1}_{t>0}\hat{F}(u, \xi) + \delta \otimes u_0$ for some modified (nonlocal) nonlinearity $\hat{F}$. Since $I$ does make perfect sense for arbitrary $b \in \mathcal{C}^{0+}$ and $f \in \mathcal{C}^{-2+}$ (which are the expected regularities of the coefficient and the right-hand side, respectively, even in the limit), this moves all the ill-defined terms into the definition of $\hat{F}$.

Verifying the equivalence is elementary as long as all functions involved are smooth: suppose that $u$ satisfies (2.5a) and apply $\partial_t - a(u)\partial_x^2$ to both sides of this equation. Denoting the expression $(\partial_t - a(u)\partial_x^2)u$ by $f$, one then has

$$\begin{aligned} f &= \hat{f} + \big((\partial_t - a(u)\partial_x^2)a(u)\big)I_1(a(u), \hat{f}) \\ &\quad - a(u)|\partial_x(a(u))|^2 I_2(a(u), \hat{f}) - 2a(u)\partial_x(a(u))I_1(a(u), \partial_x \hat{f}) \\ &= \hat{f} + a'(u) f I_1(a(u), \hat{f}) - (aa'')(u)(\partial_x u)^2 I_1(a(u), \hat{f}) \\ &\quad - (a(a')^2)(u)|\partial_x u|^2\, I_2(a(u), \hat{f}) - 2(aa')(u)\partial_x u\, I_1'(a(u), \hat{f}). \end{aligned} \tag{2.6}$$

One can rearrange the above as a fixed point equation for $\hat{f}$ by writing it as

$$\begin{aligned} \hat{f} &= (1 - a'(u)I_1(a(u), \hat{f}))f + (aa'')(u)|\partial_x u|^2 I_1(a(u), \hat{f}) \\ &\quad + (a(a')^2)(u)|\partial_x u|^2\, I_2(a(u), \hat{f}) + 2(aa')(u)\partial_x u\, I_1'(a(u), \hat{f}). \end{aligned} \tag{2.7}$$

Now we note that since $f$ is of the form $f(t, x) = \mathbf{1}_{t>0}(F(u, \xi))(t, x) + \delta(t)u_0(x)$, where $F = F(u, \xi)$ is a $\mathcal{C}^1$ function on $[0, 1] \times \mathbb{T}$, we can look for solutions to this fixed point problem that are also of the form $\hat{f}(t, x) = \mathbf{1}_{t>0}\hat{F}(t, x) + \delta(t)u_0(x)$. To see this, define the operator

$$\hat{I}_\ell^{(k)}(b, g)(z) = I_\ell^{(k)}(b, \delta \otimes g)(z) = \int \big(\partial_c^\ell P_t\big)(b(z), x - x')\partial_x^k g(x')\, dx', \tag{2.8}$$

where we use the convention $z = (t, x)$. (This is really how all terms involving $\delta$ should be interpreted in (2.5a)–(2.7).) The function $I_1(a(u), \hat{F})$ is continuous and vanishes at time 0, as does $\hat{I}_1(a(u), u_0)$ for any $u_0$ of strictly positive regularity, as one can see from (2.3) for example. Thus (2.7) can be written as a fixed point problem for $\hat{F}$:

$$\begin{aligned} \hat{F} &= \big(1 - a'(u)\, (I_1(a(u), \hat{F}) + \hat{I}_1(a(u), u_0)\big)\mathbf{1}_{t>0}F(u, \xi) \\ &\quad + (aa'')(u)|\partial_x u|^2\, (I_1(a(u), \hat{F}) + \hat{I}_1(a(u), u_0)) \\ &\quad + (a(a')^2)(u)|\partial_x u|^2\, (I_2(a(u), \hat{F}) + \hat{I}_2(a(u), u_0)) \\ &\quad + 2(aa')(u)\partial_x u\, (I_1'(a(u), \hat{F}) + \hat{I}_1'(a(u), u_0)). \end{aligned} \tag{2.5b}$$

If $u_0$ is sufficiently smooth, say $\mathcal{C}^3$, one can write the system (2.5) as a fixed point problem

$$(u, \hat{F}) = \mathcal{A}_{u_0,\xi}(u, \hat{F}), \tag{2.9}$$

where $\mathcal{A}_{u_0,\xi}$ is a contraction on a ball of $(\mathcal{C}_0^{4/3} \times \mathcal{C}_0^{-1/3})(\mathbb{T}_t)$ for small times $t$, where $\mathbb{T}_t = (-\infty, t] \times \mathbb{T}$ and $\mathcal{C}_0^\alpha$ consists of the space-time $\alpha$-Hölder regular distributions that vanish for negative times. Indeed, for the first coordinate of $\mathcal{A}_{u_0,\xi}$ this is immediate from classical Schauder estimates. For the second coordinate it suffices to notice that thanks to Remark 2.1 the right-hand side of (2.5b) is locally Lipschitz continuous from $(\mathcal{C}_0^{4/3} \times \mathcal{C}_0^{-1/3})(\mathbb{T}_t)$ to $\mathcal{C}_0^0(\mathbb{T}_t)$, which is in turn embedded into $\mathcal{C}_0^{-1/3}(\mathbb{T}_t)$, and the norm of this injection is proportional to a positive power of $t$.

Using a version of this argument with temporal weights (these allow us to deal with the possible divergence at time 0 of the various norms considered here due to the presence of the initial condition; see [7, sec. 6]), it is straightforward to show that (2.9) admits a unique local solution $(u, \hat{F})$. Furthermore, the preceding calculations show that, as long as the function $g$ given by

$$g = a'(u)(I_1(a(u), \hat{F}) + \hat{I}_1(a(u), u_0),$$

is strictly smaller than 1, one does indeed have

$$\big((\partial_t - a(u)\partial_x^2)u\big)(t, x) = (F(u, \xi))(t, x) + \delta(t)u_0(x). \tag{2.10}$$

Since, by the same reasoning as above, $g$ is continuous and $g \to 0$ as $t \to 0$, the claim follows. Moreover, if $|u_0|_{\mathcal{C}^{0+}} \le C$ for some constant $C$, then for any fixed $t_1 > 0$, the time $t_0 := \sup\{t \in [0, t_1]\colon |g(t, x)| < 1\ \forall x \in \mathbb{T}\} > 0$ can be bounded from below in terms of the $\mathcal{C}^{-2+}([-1, t_1] \times \mathbb{T})$ norm of $\hat{F}$. A reasonable solution theory of (2.9)—which of course will require a renormalization of the right-hand side of (2.9)—is expected to imply that for some $t_1 > 0$, this norm is uniformly bounded over a given family of smooth approximations of the "true" noise $\xi$, and hence $t_0$ is uniformly bounded away from 0.

It is not difficult to convince oneself that this argument is quite robust. For example, if in (2.1) the operator $\partial_x^2$ is replaced by $\partial_x^{2k}$ for some $k \in \mathbb{N}$, or if in higher dimensions $a(u)$ is matrix-valued and acts on the Hessian of $u$ in a nondegenerate way, similar arguments show the analogous equivalence, of course with $I$ built from a suitably modified family of parametrized kernels. It therefore suffices to solve equations of the type (2.5), which one can do using the framework of regularity structures as we will demonstrate in the remainder of this article.

## 3 Regularity Structures with Continuous-Parameter Dependence

It should be clear at this point that we would like to encode in our regularity structure the integration against all kernels $P(c, \cdot)$, as well as some of their derivatives with respect to $x$ and $c$. Since there is a continuum of them and since one wants to have some control over the dependence on $c$, this requires a modification of the construction in [7].

The starting point of our construction however is very similar to that given in [3, 7], and we quickly recall it here, mainly to fix notation. We fix a dimension $d \geq 1$ and on it, a scaling $\mathfrak{s} \in \mathbb{N}^d$. We assume that we are given a finite index set $\mathfrak{L} = \mathfrak{L}_+ \sqcup \mathfrak{L}_-$ as well as a map $\alpha\colon \mathfrak{L} \to \mathbb{R} \setminus \{0\}$ that is positive on $\mathfrak{L}_+$ and negative on $\mathfrak{L}_-$. We build from this a set of symbols $\mathcal{F}$ by decreeing it to be the minimal set satisfying the following properties:

- There are symbols $\Xi_i$ and $X^k$ belonging to $\mathcal{F}$ for all $i \in \mathfrak{L}_-$ and any $d$-dimensional multi-index $k$. We also write $\mathbf{1} = X^0$ and $X_i = X^{e_i}$ with $e_i$ the $i^{\text{th}}$ canonical basis vector.
- For any $\tau, \tau', \tau'' \in \mathcal{F}$, one also has $\tau\tau' \in \mathcal{F}$, and $\tau(\tau'\tau'')$ and $(\tau\tau')\tau''$ are identified. We also identify $X^k X^\ell$ with $X^{k+\ell}$, $X^k\tau$ with $\tau X^k$, and $\mathbf{1}\tau$ with $\tau$.
- For any $j \in \mathfrak{L}_+$, any $d$-dimensional multi-index $k$, and any $\tau \in \mathcal{F}$, one also has a symbol $\mathcal{I}_j^{(k)}\tau \in \mathcal{F}$.

*Remark* 3.1. It is important here that unlike [3, 7] we do *not* identify $\tau\bar{\tau}$ with $\bar{\tau}\tau$! The freedom to leave these as separate symbols will be very convenient later on.

We naturally associate degrees $|\cdot|$ to these symbols by postulating that

$$|\Xi_i| = \alpha_i, \quad |X^k| = |k|_{\mathfrak{s}}, \quad |\tau\bar{\tau}| = |\tau| + |\bar{\tau}|, \quad |\mathcal{I}_j^{(k)}\tau| = |\tau| + \alpha_j - |k|_{\mathfrak{s}}. \tag{3.1}$$

We then consider the map $\mathcal{G}\colon \mathcal{F} \to \mathcal{P}(\mathcal{F})$ (the set of all subsets of $\mathcal{F}$) defined as the minimal map ($\mathcal{P}(\mathcal{F})$ being ordered by inclusion) satisfying the following properties.

- One has $\tau \in \mathcal{G}(\tau)$ for every $\tau \in \mathcal{F}$, and one has $\mathcal{G}(\mathbf{1}) = \{\mathbf{1}\}$, $\mathcal{G}(X_i) = \{\mathbf{1}, X_i\}$, $\mathcal{G}(\Xi_i) = \{\Xi_i\}$.
- One has $\mathcal{G}(\tau\bar{\tau}) = \{\sigma\bar{\sigma}\colon \sigma \in \mathcal{G}(\tau) \text{ and } \bar{\sigma} \in \mathcal{G}(\bar{\tau})\}$.
- One has $\mathcal{G}(\mathcal{I}_\ell^{(j)}\tau) = \{\mathcal{I}_\ell^{(j)}\sigma\colon \sigma \in \mathcal{G}(\tau)\} \cup \{X^k\colon |k| < |\mathcal{I}_\ell^{(j)}\tau|\}$.

The motivation for this definition is that these properties of the set $\mathcal{G}(\tau)$ guarantee that every element of the structure group associated to our regularity structure as in [3] maps any given symbol $\tau$ into the linear span of $\mathcal{G}(\tau)$. This allows us to give the following definition of a subcritical set $\mathcal{W}$, which one should think of as any subset of $\mathcal{F}$ that generates an actual regularity structure (one in which the set of possible degrees is locally finite and bounded from below).

DEFINITION 3.2. A subset $\mathcal{W} \subset \mathcal{F}$ is said to be *subcritical* if it satisfies the following properties:

- If $\tau \in \mathcal{W}$, then $\mathcal{G}(\tau) \subset \mathcal{W}$.
- For every $\gamma \in \mathbb{R}$, the set $\{\tau \in \mathcal{W}\colon |\tau| < \gamma\}$ is finite.

It is said to be *normal* if, whenever $\tau\bar{\tau} \in \mathcal{W}$, one has $\{\tau, \bar{\tau}\} \subset \mathcal{W}$ and, whenever $\mathcal{I}_\ell^{(j)}\tau \in \mathcal{W}$, one has $\tau \in \mathcal{W}$.

As shown in [3, 7], every locally subcritical stochastic PDE (or system thereof) naturally determines a normal subcritical set $\mathcal{W}$. From now on, we consider $\mathcal{W}$ to be fixed, and we will only consider elements $\tau \in \mathcal{W}$.

### 3.1 A Regularity Structure

In [3, 7], one then constructs a regularity structure by taking the vector space $\langle\mathcal{W}\rangle$ generated by $\mathcal{W}$ as the structure space (graded by the notion of degree given in (3.1)) and endowing it with a suitable structure group. In our situation, to encode parameter dependence, we instead assign to each element of $\mathcal{W}$ a typically infinite-dimensional subspace of the structure space. In order to encode this, we first define the number of parameters $[\tau]$ in a symbol $\tau$ recursively by setting

$$[X^k] = [\Xi_i] = 0, \quad [\tau\bar{\tau}] = [\tau] + [\bar{\tau}], \quad [\mathcal{I}_j^{(k)}\tau] = [\tau] + 1.$$

*Remark* 3.3. One could in principle encode some parametrization of the noises as well, by setting $[\Xi_i] = 1$, or we could even allow the number of parameters to depend on the element of $\mathfrak{L}$ we consider. Since this generality is not used in what follows, we refrain from doing so here.

We also assume that we are given a real Banach space $\mathcal{B}$, and we write $\mathcal{B}_k$ for the $k$-fold tensor product of $\mathcal{B}$ with itself, completed under the projective cross norm. In particular, we have a canonical dense embedding of $\mathcal{B}_k \otimes \mathcal{B}_\ell$ into $\mathcal{B}_{k+\ell}$. We also use the convention $\mathcal{B}_0 = \mathbb{R}$. Given a normal subcritical set of symbols $\mathcal{W}$, we then construct a regularity structure from it in such a way that each symbol $\tau \in \mathcal{W}$ determines an infinite-dimensional subspace $T_\tau$ of the structure space $T$, isometric to $\mathcal{B}_{[\tau]}$. To wit, we set

$$T = \bigoplus_\alpha T_\alpha, \quad T_\alpha := \bigoplus_{|\tau|=\alpha} T_\tau, \quad T_\tau := \mathcal{B}_{[\tau]} \otimes \langle\tau\rangle, \tag{3.2}$$

and equip the spaces $T_\alpha$ with their natural norms. Here and below, we write $\langle A\rangle$ for the free real vector space generated by a set $A$, and we make a slight abuse of notation by also writing $\langle\tau\rangle$ as a shorthand for $\langle\{\tau\}\rangle$ when $\tau$ is a symbol in $\mathcal{W}$. (By the definition of subcriticality, there are only finitely many symbols $\tau$ with $|\tau| = \alpha$, so that the $T_\alpha$ are naturally endowed with a Banach space structure. This is not the case for $T$ itself though, but we view it as a topological vector space in the usual way.) For $\tau$ with $[\tau] = 0$, we also identify $T_\tau$ with $\langle\tau\rangle$.

The space $T$ comes equipped with a number of natural operations. For every $i \in \{1, \dots, d\}$, we have an abstract differentiation $D_i$ acting on $\langle\mathcal{F}\rangle$ by setting $D_i X_j = \delta_{ij}\mathbf{1}$, $D_i\mathbf{1} = 0$, $D_i\mathcal{I}_j^{(k)}(\tau) = \mathcal{I}_j^{(k+e_i)}(\tau)$, and then extending it to all other symbols by enforcing that Leibniz's rule holds. For any $\tau \in \mathcal{W}$ such that $D_i\tau \in \langle\mathcal{W}\rangle$ and any $b \in \mathcal{B}_{[\tau]}$, we then set

$$\mathscr{D}_i(b \otimes \tau) = b \otimes D_i\tau.$$

This is indeed well-defined since $D_i\tau$ is a linear combination of elements $\sigma$ with $[\sigma] = [\tau]$. Similarly, for the abstract product, whenever $\tau, \overline{\tau}, \tau\overline{\tau} \in \mathcal{W}$, $b \in \mathcal{B}_{[\tau]}$, $\overline{b} \in \mathcal{B}_{[\overline{\tau}]}$, we set

$$(b \otimes \tau)(\overline{b} \otimes \overline{\tau}) = (b \otimes \overline{b}) \otimes \tau\overline{\tau},$$

with $b \otimes \overline{b}$ interpreted as an element of $\mathcal{B}_{[\tau\overline{\tau}]}$. (Here it is convenient that $\tau\overline{\tau}$ and $\overline{\tau}\tau$ aren't identified since it avoids being forced to deal with symmetric tensor products.) Finally, we have a large number of abstract integration operators: for any $j \in \mathfrak{L}_+$ and any $b \in \mathcal{B}$, a map $\mathcal{I}_j^{(k),b}$ is defined as the linear extension of

$$\mathcal{I}_j^{(k),b}(\overline{b} \otimes \tau) := (b \otimes \overline{b}) \otimes \mathcal{I}_j^{(k)}\tau,$$

defined on those $T_\tau$ for which $\mathcal{I}_j^{(k)}\tau \in \mathcal{W}$.

So far we have not addressed the structure group at all, but its inductive construction as in [7, thm. 5.14] is virtually identical in our setting. More precisely, as in [7, defs. 4.6 and 5.25] the group $G$ consists of those continuous linear operators $\Gamma\colon T \to T$ satisfying the following properties:

- For any $\alpha \in \mathbb{R}$, one has $(\Gamma - \mathrm{id})\colon T_\alpha \to T_{<\alpha}$.
- One has $\Gamma\mathbf{1} = \mathbf{1}$, $\Gamma\Xi_i = \Xi_i$, and there are constants $c_i$ such that $\Gamma X_i = X_i - c_i\mathbf{1}$.
- For any two symbols $\tau, \overline{\tau}$ in $\mathcal{W}$ such that $\tau\overline{\tau} \in \mathcal{W}$ and any $a \in T_\tau$, $\overline{a} \in T_{\overline{\tau}}$, one has $\Gamma(a\overline{a}) = (\Gamma a)(\Gamma\overline{a})$.
- For any $\tau \in \mathcal{W}$ such that $D_i\tau \in \langle\mathcal{W}\rangle$ and any $a \in T_\tau$, one has $\Gamma\mathscr{D}_i a = \mathscr{D}_i\Gamma a$.
- For any $\tau \in \mathcal{W}$ such that $\mathcal{I}_\ell^{(k)}\tau \in \mathcal{W}$, any $a \in T_\tau$, and any $b \in \mathcal{B}$, one has

$$\big(\Gamma\mathcal{I}_\ell^{(k),b} - \mathcal{I}_\ell^{(k),b}\Gamma\big)a \in \langle\{X^k\colon k \in \mathbb{N}^d\}\rangle.$$

As in [7], one can show that this is indeed a group. The definitions of $G$ and $\mathcal{G}$ also guarantee that, for any $\tau \in \mathcal{W}$, any $\Gamma \in G$ maps $T_\tau$ to $\bigoplus_{\sigma\in\mathcal{G}(\tau)} T_\sigma$, which is indeed a subspace of $T$ by the assumption on $\mathcal{W}$.

From now on, we write $\mathscr{T}$ for the regularity structure with structure space $T$ and structure group $G$ given as above with the specific choice

$$\mathcal{B} = \mathcal{C}^{-N}(\mathfrak{K}) \tag{3.3}$$

for some compact parameter space $\mathfrak{K}$, which is assumed for simplicity to be a subset of a $\mathbb{R}^{d_1}$, as well as some sufficiently large $N > 0$ to be determined later.

### 3.2 Admissible Models

We assume henceforth that $\mathfrak{L}_+$ and $\mathfrak{L}_-$ are singletons, and therefore omit the lower indices in $\Xi$, $\mathcal{I}$. This is purely for the sake of notational convenience; this section immediately extends to the general case. We also omit $k$ in $\mathcal{I}^{(k)}$ and $\mathcal{I}^{(k),b}$ if $k = 0$, and we set $\alpha = |\Xi|$ and $\beta = |\mathcal{I}\tau| - |\tau|$.

*Assumption* 3.4. We are given a family of kernels $(K^{(c)})_{c\in\mathfrak{K}}$, which, along with their derivatives with respect to $c$ up to any finite order, are uniformly compactly supported and $\beta$-smoothing in the sense of [7, ass. 5.1].

*Remark* 3.5. Think of $K^{(c)}$ as the heat kernel with parameter $c$ as in (2.3). These are of course not compactly supported and do not satisfy [7, ass. 5.4]. However, it is always possible to choose kernels $K^{(c)}$ satisfying these properties and such that their projection onto $\Lambda = [-1, \infty) \times \mathbb{T}^{d-1}$ (obtained by adding all integer translates) agrees with the heat kernel, so that if $\eta$ is a distribution on $\mathbb{R}_+ \times \mathbb{T}^{d-1}$, one has $K^{(c)} * \eta = P(c, \cdot) * \eta$ on $[0, 1] \times \mathbb{T}^{d-1}$.

An example of explicit construction of $K^{(c)}$ goes as follows. We fix a smooth function $\chi\colon \mathbb{R}^{d-1} \to \mathbb{R}$ supported in $[-\frac{3}{4}, \frac{3}{4}]^{d-1}$ such that $\sum_{n\in\mathbb{Z}^{d-1}} \chi(x+n) = 1$ for all $x$ and such that $\chi(x) = 1$ for $x \in [-\frac{1}{4}, \frac{1}{4}]^{d-1}$. We also fix a smooth cutoff function $\tilde{\chi}\colon \mathbb{R} \to \mathbb{R}$ such that $\tilde{\chi}(t) = 1$ for $t \geq -1$ and $\tilde{\chi}(t) = 0$ for $t \leq -2$. Finally, we fix a finite collection of smooth functions $\phi_k\colon \mathbb{R}^d \to \mathbb{R}$ that are supported on $[-2, -1] \times [-1, 1]^{d-1}$ and such that $\int \phi_k(z) z^\ell \, dz = \delta_{k\ell}$ for all multi-indices $k, \ell$ with $|k| \vee |\ell| \leq r$, with $r$ as in [7, ass. 5.4]. Then, it is an easy exercise to check that the choice of $K^{(c)}$ given by

$$K^{(c)}(z) = \tilde{K}^{(c)}(z) - \sum_{|k|\leq r} \phi_k(z) \int \tilde{z}^k \tilde{K}^{(c)}(\tilde{z}) d\tilde{z},$$
$$\tilde{K}^{(c)}(t, x) = \sum_{n\in\mathbb{Z}^{d-1}} P(c, t, x + n) \chi(x) \tilde{\chi}(t),$$

does satisfy assumptions 5.1 and 5.4 of [7] and furthermore behaves as advertised.

With this construction, one also has

$$(\partial_t - c\Delta) K^{(c)} = \delta_0 + f^{(c)}, \tag{3.4}$$

where $\delta_0$ is the Dirac mass at the space-time origin and $f^{(c)}$ are compactly (and away from the origin) supported smooth functions, depending smoothly on $c$.

Furthermore, for any distribution $\zeta$ on $\mathfrak{K}$, any $c \in \mathfrak{K}$, and any $\ell \in \mathbb{N}^{d_1}$, we write

$$K^\zeta := \zeta(K^{(\cdot)}), \quad K^{c;\ell} := K^{\partial^\ell \delta_c}.$$

By the assumption on $K$, $K^\zeta$ is also $\beta$-smoothing in the sense of [7, ass. 5.1] and, when considering its decomposition $K^\zeta = \sum_{n\geq 0} K_n^\zeta$, one has a bound of the type $|D^k K_n^\zeta| \lesssim 2^{n(|\mathfrak{s}|-\beta+|k|_\mathfrak{s})} |\zeta|_{\mathcal{C}^{-N_0}}$ for any fixed $N_0 > 0$.

As our notation suggests, we want the maps $\mathcal{I}^\zeta$ to correspond to integrations against the kernels $K^\zeta$, which is encoded in the definition of admissibility in the present setting.

DEFINITION 3.6. In the above setting, a model $(\Pi, \Gamma)$ is admissible for $\mathscr{T}$ if, for all $\alpha \in A$, $\tau \in \mathcal{W}$, such that $|\tau| = \alpha$ and $\mathcal{I}\tau \in \mathcal{W}$, for all $\sigma \in T_\tau$, $\zeta \in \mathcal{C}^{-N}$,

$x \in \mathbb{R}^d$, and $\varphi \in C_0^\infty$, the following identity holds:

$$(\Pi_x \mathcal{I}^\xi \sigma)(\varphi) = (K^\xi * \Pi_x \sigma)(\varphi) - \int \sum_{|k|_\mathfrak{s} \le \alpha+\beta} \frac{(y-x)^k}{k!} \Pi_x \sigma(D^k K^\xi(x-\cdot))\varphi(y)dy.$$

One can define the maps $\mathcal{J}^\xi$ as in [7], with $K$ therein replaced by $K^\xi$. With this notation the second term on the right-hand side above can also be written as $(\Pi_x \mathcal{J}^\xi(x)\sigma)(\varphi)$.

In the following we borrow the notation $\psi_x^\lambda$, $\|\Pi - \Pi'\|_{\gamma;B}$, $\|\Gamma - \Gamma'\|_{\gamma;B}$ from [7], and denote by $\mathscr{B}$ the set $\mathcal{B}_{-\lfloor\alpha\rfloor}$ of test functions considered there. (This is in order to prevent confusion with the scale of spaces $\mathcal{B}_k$.) In fact, the lower indices in the norms of the models will usually be omitted for brevity, since the dependence on them does not play any role in our discussion.

### 3.3 Constructing Models

In principle, if one has a sufficiently robust way of building a model (or a family of models with some continuity properties) for the (usual) regularity structure determined by $\mathcal{W}$, one can also build an admissible model for the parametrized regularity structure $\mathscr{T}$. Such a robust way of building models is developed in great generality in [4]. To be self-contained regarding the assumptions required to recall some of its results, we restrict our attention to the Gaussian case and refer the reader to [4] for more general noises that fit in the theory.

*Assumption* 3.7. Suppose we are given a centered, Gaussian, translation invariant, $\mathcal{S}'(\mathbb{R}^d)$-valued random variable $\xi$ such that there exists a distribution $\mathscr{C}$ whose singular support is contained in $\{0\}$, which satisfies

$$\mathbf{E}(\xi(f)\xi(g)) = \mathscr{C}\left(\int f(z-\cdot)g(z)dz\right)$$

for all test functions $f, g \in \mathcal{S}(\mathbb{R}^d)$. Writing $z \mapsto \mathscr{C}(z)$ for the smooth function that determines $\mathscr{C}$ away from 0, it is furthermore assumed that any test function $g$ satisfying $D^k g(0) = 0$ for all multi-index $k$ with $|k|_\mathfrak{s} < -|\mathfrak{s}| - 2\alpha$, one has

$$\mathscr{C}(g) = \int \mathscr{C}(z)g(z)dz.$$

Finally, there exists a $\kappa > 0$ such that for all multi-index $k$

$$\sup_{0<|z|_\mathfrak{s}\le 1} |D^k \mathscr{C}(z)|\,|z|_\mathfrak{s}^{|k|_\mathfrak{s}-2\alpha-\kappa} < \infty.$$

The final assumption on $\mathcal{W}$ is what is referred to as super-regularity in [4], which in the present setting reads as follows. Define, similarly to $[\cdot]$, the number of noises $[\![\tau]\!]$ in a symbol $\tau$ recursively by

$$[\![X^k]\!] = 0, \quad [\![\Xi]\!] = 1, \quad [\![\tau\bar\tau]\!] = [\![\tau]\!] + [\![\bar\tau]\!], \quad [\![\mathcal{I}^{(k)}\tau]\!] = [\![\tau]\!].$$

*Assumption* 3.8. All $\tau \in \mathcal{W}$ with $[\![\tau]\!] \geq 2$ satisfy $|\tau| > \alpha$ and $|\tau| > -|\mathfrak{s}|/2$. If $[\![\tau]\!] \geq 3$, then also $|\tau| > -(|\mathfrak{s}| + \alpha)$, while if $[\![\tau]\!] = 2$, then also $|\tau| > -2(|\mathfrak{s}| + \alpha)$ holds.

Take, as in the introduction, a compactly supported, nonnegative, symmetric smooth function $\rho$ integrating to 1, and set $\rho^\varepsilon(t,x) = \varepsilon^{-|\mathfrak{s}|}\rho(z_1\varepsilon^{-\mathfrak{s}_1}, \dots, z_d\varepsilon^{-\mathfrak{s}_d})$. Under the above assumptions, we wish to construct a family of admissible models $(\widehat{\Pi}^\varepsilon, \widehat{\Gamma}^\varepsilon)_{\varepsilon \in [0,1]}$ that is continuous in a suitable sense in the $\varepsilon \to 0$ limit, and which satisfy $\widehat{\Pi}^\varepsilon_z \Xi = \rho^\varepsilon * \xi$ (here and below we use the natural convention of $\rho^0 *$ denoting the identity).

Recalling (3.3) and setting $N_0 = N + d_1 + 1$, given a finite set $\widetilde{B} \subset \mathcal{C}^{-N_0}$, let $S_{\widetilde{B}}$ be the set of simple elements of the form

$$a = \left(\bigotimes_{i=1}^{[\tau]} \zeta_i\right) \otimes \tau \quad \text{with } \tau \in \mathcal{W},\ \zeta_i \in \widetilde{B},$$

and let $S = \bigcup_{\widetilde{B}} S_{\widetilde{B}}$ (notice that $S \not\subseteq T$ since one has $N_0 > N$). Any $a \in S_{\widetilde{B}}$ can be mapped to an element $\iota(a)$ of the structure space $T_{\widetilde{B}}$ for the regularity structure $\mathscr{T}_{\widetilde{B}}$ built from $\mathfrak{L}_+ = \widetilde{B}$, $\alpha(\mathfrak{L}_+) = \{\beta\}$, by setting recursively

$$\iota(\Xi) = \Xi, \quad \iota(X^k) = X^k, \quad \iota(a\bar{a}) = \iota(a)\iota(\bar{a}), \quad \iota(\mathcal{I}^{(\ell),\zeta}a) = \mathcal{I}^{(\ell)}_\zeta \iota(a).$$

Let $Z^\varepsilon = (\Pi^\varepsilon, \Gamma^\varepsilon)$ for $\varepsilon \in [0,1]$ be the family of BPHZ models for $\mathscr{T}_{\widetilde{B}}$ as constructed in [3, 4], which satisfy in particular $\Pi^\varepsilon_z \Xi = \rho^\varepsilon * \xi$. One can then define the random distributions

$$\overline{\Pi}^\varepsilon_x a := \Pi^\varepsilon_x \iota(a), \tag{3.5}$$

for $a \in S$. Note that formally the right-hand side also depends on $\widetilde{B}$ (the regularity structure $\mathscr{T}_{\widetilde{B}}$ in which $\iota(a)$ takes values depends on it, as well as the model $Z^\varepsilon$), but our construction is such that different choices of $\widetilde{B}$ yield the same right hand side in (3.5). By [4], the random fields $\overline{\Pi}_x$ satisfy the bounds

$$\begin{aligned} \sup_{0 \neq a \in S} |a|^{-p} \mathbf{E} \sup_{x,\lambda,\psi} \lambda^{-p|\tau|} \big|\big(\overline{\Pi}^\varepsilon_x a\big)\big(\psi^\lambda_x\big)\big|^p &\lesssim 1, \\ \sup_{0 \neq a \in S} |a|^{-p} \mathbf{E} \sup_{x,\lambda,\psi} \lambda^{-p|\tau|} \big|\big(\big(\overline{\Pi}^\varepsilon_x - \overline{\Pi}^0_x\big)a\big)\big(\psi^\lambda_x\big)\big|^p &\lesssim \varepsilon^{p\theta}, \end{aligned} \tag{3.6}$$

with some $\theta > 0$, where here and below the second supremum is taken over $x$ in some compact set, $\lambda \in (0,1]$, and $\psi \in \mathscr{B}$. The random field $\overline{\Pi}$ can then be turned into an admissible model for $\mathscr{T}$ in the following sense.

THEOREM 3.9. *There exist admissible models $\widehat{Z}^\varepsilon = (\widehat{\Pi}^\varepsilon, \widehat{\Gamma}^\varepsilon)$ with $\varepsilon \in [0,1]$ for $\mathscr{T}$ such that for all $a \in S \cap T$, $\widehat{\Pi}^\varepsilon_x a = \overline{\Pi}^\varepsilon_x a$ almost surely, and that one has the bounds*

$$\mathbf{E}(\|\widehat{\Pi}^\varepsilon\| + \|\widehat{\Gamma}^\varepsilon\|)^p \lesssim 1,$$
$$\mathbf{E}(\|\widehat{\Pi}^\varepsilon - \widehat{\Pi}^0\| + \|\widehat{\Gamma}^\varepsilon - \widehat{\Gamma}^0\|)^p \lesssim \varepsilon^{p\theta}.$$

PROOF. Define the set $S' \subset S$ similarly to $S$, but with $\mathcal{C}^{-N_0}$ replaced by $\{\partial^\ell \delta_c \colon c \in \mathfrak{K}, |\ell| \le N\} \subset \mathcal{C}^{-N_0}$. For $\tau \in \mathcal{W}$, $c \in \mathfrak{K}^{[\tau]}$, and $\ell \in (\mathbb{N}^{d_1})^{[\tau]}$ with $|\ell_i| \le N$, denote $a_{c,\ell}(\tau) = (\bigotimes_{i=1}^{[\tau]} \partial^{\ell_i} \delta_{c_i}) \otimes \tau$. From (3.6), we have

$$\mathbf{E} \sup_{x,\lambda,\psi} \lambda^{-p|\tau|} \big| \big(\overline{\Pi}^\varepsilon_x (a_{c,\ell}(\tau) - a_{\overline{c},\ell}(\tau)\big)(\psi^\lambda_x)\big|^p \lesssim |c - \overline{c}|^p$$

for any $c, \overline{c} \in \mathfrak{K}^{[\tau]}$. Choosing $p$ large enough, by Kolmogorov's continuity theorem one has a continuous modification $(\widehat{\Pi}_x a_{c,\ell}(\tau))_{c \in \mathfrak{K}^{[\tau]}}$ such that the admissibility condition is satisfied almost surely, and that one has the bound

$$\mathbf{E} \sup_{x,\lambda,\psi} \sup_{c \in \mathfrak{K}^{[\tau]}} \lambda^{-p|\tau|} \big| \big(\widehat{\Pi}^\varepsilon_x a_{c,\ell}(\tau)\big)\big(\psi^\lambda_x\big)\big|^p \lesssim 1.$$

Note that a generic element of $S \cap T$ is of the form $a = (\bigotimes_{i=1}^{[\tau]} \zeta_i) \otimes \tau$, with $\zeta_i \in \mathcal{C}^{-N} = \mathcal{B}$. Hence on $S \cap T$ we can define

$$\widehat{\Pi}^\varepsilon_x a := \big(\textstyle\bigotimes_{i=1}^{[\tau]} \zeta_i\big)\big(c \mapsto \widehat{\Pi}^\varepsilon_x a_{c,0}(\tau)\big),$$

and extending these maps to all of $T$ by linearity and continuity, we get maps $\widehat{\Pi}^\varepsilon_x$ that are admissible and that satisfy

$$\mathbf{E} \sup_{x,\lambda,\psi} \sup_{0 \ne a \in T} |a|^{-p} \lambda^{-p|\tau|} \big| \big(\widehat{\Pi}^\varepsilon_x a\big)\big(\psi^\lambda_x\big)\big|^p \lesssim 1.$$

The corresponding bounds on the differences $\widehat{\Pi}^\varepsilon_x - \widehat{\Pi}^0_x$ are obtained similarly, so it remains to treat the maps $\widehat{\Gamma}^\varepsilon$.

We proceed inductively. The definition of, and the appropriate bounds on, $\widehat{\Gamma}^\varepsilon_{xy}\tau$ if $\tau = \Xi$ or $X^k$, are trivial. Given $\widehat{\Gamma}^\varepsilon_{xy}(\zeta \otimes \tau)$ and $\widehat{\Gamma}^\varepsilon_{xy}(\overline{\zeta} \otimes \overline{\tau})$ with the right bounds, we set

$$\widehat{\Gamma}_{xy}((\zeta \otimes \overline{\zeta}) \otimes \tau\overline{\tau}) = \big(\widehat{\Gamma}^\varepsilon_{xy}(\zeta \otimes \tau)\big)\big(\widehat{\Gamma}^\varepsilon_{xy}(\overline{\zeta} \otimes \overline{\tau})\big),$$

which one can bound by

$$\begin{aligned} \big\|\widehat{\Gamma}_{xy}((\zeta \otimes \overline{\zeta}) \otimes \tau\overline{\tau})\big\|_m &\lesssim \sum_{m_1+m_2=m} \|x-y\|_{\mathfrak{s}}^{|\tau|-m_1} |\zeta|_{\mathcal{B}_{[\tau]}} \|x-y\|_{\mathfrak{s}}^{|\overline{\tau}|-m_2} |\zeta|_{\mathcal{B}_{[\overline{\tau}]}} \\ &\lesssim \|x-y\|_{\mathfrak{s}}^{|\tau\overline{\tau}|-m} |\zeta \otimes \overline{\zeta}|_{\mathcal{B}_{[\tau\overline{\tau}]}}, \end{aligned}$$

where we used our assumption on the spaces $\mathcal{B}_k$ to obtain the second line. Given $\widehat{\Gamma}^\varepsilon_{xy}\zeta \otimes \tau$, we also set $\widehat{\Gamma}^\varepsilon_{xy}(\zeta \otimes \mathscr{D}^i \tau) = \mathscr{D}_i \widehat{\Gamma}^\varepsilon_{xy}(\zeta \otimes \tau)$, for which the correct bounds follow automatically.

The only step to finish the induction is thus to define and bound $\hat{\Gamma}^{\varepsilon}_{xy}\zeta \otimes \mathcal{I}\tau$, provided $\hat{\Gamma}^{\varepsilon}_{xy}\zeta' \otimes \tau$ are known. This is done as in [7, thm. 5.14]: for $\zeta_1 \in \mathcal{B}$, $a = \zeta' \otimes \tau$, we set

$$\hat{\Gamma}^{\varepsilon}_{xy}\mathcal{I}^{\zeta_1}a = \mathcal{I}^{\zeta_1}a + \mathcal{I}^{\zeta_1}\big(\hat{\Gamma}^{\varepsilon}_{xy}a - a\big) + \big(\mathcal{J}^{\zeta_1}(x)\hat{\Gamma}^{\varepsilon}_{xy}a - \hat{\Gamma}^{\varepsilon}_{xy}\mathcal{J}^{\zeta_1}(y)a\big). \tag{3.7}$$

The first term on the right-hand side is harmless. Bounding the second one is immediate:

$$\big\|\mathcal{I}^{\zeta_1}\big(\hat{\Gamma}^{\varepsilon}_{xy}a - a\big)\big\|_m \lesssim |\zeta_1|_{\mathcal{B}}\big\|\hat{\Gamma}^{\varepsilon}_{xy}a - a\big\|_{m-\beta} \lesssim |\zeta_1|_{\mathcal{B}}|\zeta'|_{\mathcal{B}_{[\tau]}}\|x-y\|^{|\tau|+\beta-m}$$

thanks to the assumed bound on $\hat{\Gamma}^{\varepsilon}_{xy}\zeta' \otimes \tau$. Using again the assumptions on the spaces $\mathcal{B}_k$, this is precisely the required bound. To bound the third term on the right-hand side of (3.7), it suffices to recall [7, lem. 5.21], with the kernel $K$ therein replaced by $K^{\zeta_1}$. Having the required bounds on elements of the form $\hat{\Gamma}^{\varepsilon}_{xy}(\zeta_1 \otimes \zeta' \otimes \mathcal{I}\tau)$, one can extend $\hat{\Gamma}^{\varepsilon}_{xy}$ to all $a \in T_{\mathcal{I}\tau}$ once again via linearity and continuity. It is straightforward to check that the above-defined maps $\hat{\Gamma}^{\varepsilon}_{xy}$ do indeed belong to $G$, and hence the proof is finished. □

*Remark* 3.10. Let us comment briefly on the renormalization procedure implicit in the construction (3.5). In the standard situation considered in [3, 4], the BPHZ renormalization procedure assigns to each symbol $\tau \in \mathcal{W}$ with $\tau \neq \mathbf{1}$ and $|\tau| \leq 0$ a constant $C^{\varepsilon}_{\tau}$. (In the notation of [3, eq. 6.22], one has $C^{\varepsilon}_{\tau} = g_{-}(\mathbf{\Pi}^{\varepsilon})(\iota_{\circ}\tau) = \mathbf{E}(\mathbf{\Pi}^{\varepsilon}\tau)(0)$, with $\mathbf{\Pi}^{\varepsilon}$ the canonical lift of the mollified noises.) This choice then allows us to define a renormalized model by [3, thms 6.17, 6.27] that was shown in [4, thms 2.14, 2.30] to enjoy very strong stability properties.

The construction given above is essentially the same, but now each symbol $\tau$ determines a smooth *function* $C^{\varepsilon}_{\tau}\colon \mathfrak{K}^{[\tau]} \to \mathbb{R}$, where

$$C^{\varepsilon}_{\tau}(c) = \mathbf{E}\mathbf{\Pi}^{\varepsilon}\bigg(\bigg(\bigotimes_{i=1}^{[\tau]}\delta_{c_i}\bigg) \otimes \tau\bigg)(0). \tag{3.8}$$

The construction of the renormalized model is then the same as in [3].

## 4 Lifting the Operator *I*

We continue within the setting of the previous section. Given now that we have abstract integration operators $\mathcal{I}^{\zeta}$ on $T$ that that can in principle be used as in [7, sec. 4] to build the operation of convolution with any of the $K^{\zeta}$, we are also able to construct the abstract counterpart of the operators $I^{(\ell)}_k$, acting on suitable spaces of modeled distributions.

From now on we assume $d > 1$, and the first coordinate will be viewed as time. We work with $\mathcal{D}^{\gamma,\eta}_P$ spaces defined as in [7, sec. 6], with $P = \{(0,x)\colon x \in \mathbb{T}^{d-1}\}$. It will be clear that apart from notational inconvenience there is no fundamental obstacle to obtaining analogous results for more complicated weighted spaces like those considered in [6] that are suitable for solving initial boundary value problems.

Given the setup of the previous section and an admissible model $(\Pi, \Gamma)$, one can define the maps $\mathcal{K}_m^\zeta$ by replacing $\mathcal{I}$ and $K$ in [7, eq. 5.15] by $\mathcal{I}^{(m),\zeta}$ and $D_m K^\zeta$, respectively, provided $|m|_\mathfrak{s} < \beta$. As before, we denote $\mathcal{K}_m^{c;\ell} := \mathcal{K}_m^{\partial^\ell \delta_c}$, and for $m = 0$ the lower index is omitted.

We now define the lift of $I$ by a sort of higher-order freezing of coefficients where, around a given fixed point $z_0$, we don't just describe $I(b, f)$ by $\mathcal{K}^{b(z_0);0} f$, but also use higher-order information about $b$. Set, with $\overline{b} = \langle b, \mathbf{1}\rangle$ and $\widehat{b} = b - \overline{b}$,

$$\mathfrak{I}_k^{(m)}(b, f)(z) := \sum_{|\ell| \le N'} \frac{(\widehat{b}(z))^\ell}{\ell!} \big(\mathcal{K}_m^{\overline{b}(z);k+\ell} f\big)(z), \tag{4.1}$$

where $N'$ is a sufficiently large integer. (How large exactly will be specified in the statement of Theorem 4.4 below. Since the exact value of $N'$ does not make much of a difference for our purpose, we do not explicitly keep track of it in our notation.) In the following we treat only $\mathfrak{I} := \mathfrak{I}_0^{(0)}$. The Schauder estimate for $\mathfrak{I}_k^{(m)}$ can then be formally obtained by changing the family of kernels $(K^{(c)})_{c\in\mathfrak{K}}$ to $(\partial_c^k \partial_x^m K^{(c)})_{c\in\mathfrak{K}}$, as well as $\beta$ to $\beta - |m|_\mathfrak{s}$, and apply the Schauder estimate for the map $\mathfrak{I}$ built from this family.

Note that the definition (4.1) is very reminiscent of how one composes modeled distributions with smooth functions $F$; see [7, sec. 4.2]. To justify this analogy, one needs a substitute for the Taylor expansion of $F$, which is precisely the content of Corollary 4.3 below. Thanks to this (of course not coincidental; see Remark 4.6 below) similarity, the Schauder estimates for $\mathfrak{I}$ will follow immediately from the one for constant coefficients (Theorem 4.2 below), and a straightforward adaptation of the proof of [7, prop. 6.13].

Recall that we previously fixed $\alpha \in \mathbb{R}$, $\beta > 0$. Fix further some $\gamma_1, \gamma_2, \overline{\gamma} > 0$ and $\eta_1$, $\eta_2$, and $\overline{\eta}$ such that

$$\overline{\gamma} \le \big(\gamma_1 + (\alpha+\beta)\wedge 0\big) \wedge \big(\gamma_2 + \beta\big), \quad \alpha > \eta_2 > -\mathfrak{s}_1,$$
$$\eta_2 + \beta > \overline{\eta}, \quad \eta_1 \ge \overline{\eta} \vee 0.$$

*Remark* 4.1. Note that if $\alpha + \beta \ge 0$ and $\eta_1 \ge 0$, then one can simply choose $\overline{\gamma} = \gamma_1 = \gamma_2 + \beta$ and $\overline{\eta} = \eta_1 < \eta_2 + \beta < \alpha + \beta$.

We assume henceforth that the kernels $K^{(c)}$ are nonanticipative, namely that they vanish for negative times. One then has the following (see [7, thm. 7.1]):

THEOREM 4.2. *With $\kappa = (\eta_2 + \beta - \overline{\eta})/\mathfrak{s}_1$, for any $\zeta \in \mathcal{C}^{-N}$, $f \in \mathcal{D}_P^{\gamma_2,\eta_2}$, and any $t \in (0, 1]$, one has*

$$|||\mathcal{K}^\zeta \mathbf{R}^+ f|||_{\gamma_2+\beta,\overline{\eta};t} \lesssim t^\kappa |\zeta|_{\mathcal{C}^{-N}} |||f|||_{\gamma_2,\eta_2;t}. \tag{4.2}$$

COROLLARY 4.3. *Let* $f \in \mathcal{D}_P^{\gamma_2,\eta_2}(V)$. *Then for* $c, \bar{c} \in \mathfrak{K}$, $\ell \in \mathbb{N}^{d_1}$, $m \geq 0$ *with* $m + |\ell| + d_1 + 1 \leq N$, *and any* $t \in (0,1]$, *one has*

$$\Big\| \mathcal{K}^{c;\ell}\mathbf{R}^+ f - \sum_{|k|\leq m} \frac{(c-\bar{c})^k}{k!} \mathcal{K}^{\bar{c};\ell+k}\mathbf{R}^+ f \Big\|_{\gamma_2+\beta,\bar{\eta};t} \lesssim |c-\bar{c}|^{m+1} t^\kappa \|\!|f|\!\|_{\gamma_2,\eta_2;t}. \tag{4.3}$$

PROOF. Simply apply Theorem 4.2 with $\zeta = \partial^\ell \delta_c - \sum_{|k|\leq m} \frac{(c-\bar{c})^k}{k!} \partial^{k+\ell}\delta_{\bar{c}}$. □

THEOREM 4.4. *Assume the above setting and suppose that* $b \in \mathcal{D}_P^{\gamma_1,\eta_1}(V)$ *is* $\mathfrak{K}$*-valued,* $f \in \mathcal{D}_P^{\gamma_2,\eta_2}$, *where* $V$ *is a functionlike sector with lowest nonzero homogeneity* $\alpha_1$ *and* $N'\alpha_1 > \gamma_2 + \beta$. *Then* $\mathfrak{I}(b,f) \in \mathcal{D}_P^{\bar{\gamma},\bar{\eta}}$.

*If* $\tilde{b} \in \mathcal{D}_P^{\gamma_1,\eta_1}(V,\tilde{\Gamma})$ *and* $\tilde{f} \in \mathcal{D}_P^{\gamma_2,\eta_2}(T,\tilde{\Gamma})$ *with another admissible model* $(\tilde{\Pi},\tilde{\Gamma})$, *then one has the following bound for any* $t \in (0,1]$:

$$\|\!|\mathfrak{I}(b,\mathbf{R}^+ f);\mathfrak{I}(\tilde{b},\mathbf{R}^+\tilde{f})|\!\|_{\bar{\gamma},\bar{\eta};t} \lesssim t^\kappa\big(\|\!|b;\tilde{b}|\!\|_{\gamma_1,\eta_1;t} + \|\!|f;\tilde{f}|\!\|_{\gamma_2,\eta_2;t} + \|(\Pi,\Gamma)-(\tilde{\Pi},\tilde{\Gamma})\|_{\bar{\gamma};2}\big).$$

*Moreover, if* $\alpha+\beta > 0$, *then the identity*

$$\mathcal{R}\mathfrak{I}(b,f) = I(\mathcal{R}b,\mathcal{R}f) \tag{4.4}$$

*holds, where* $I$ *is defined as in* (2.4).

*Remark* 4.5. Note that if $\alpha_1 + \alpha + \beta < 0$, then the equality (4.4) fails to hold in general even for canonical models built from a smooth noise.

PROOF. Denoting $F^{(\ell)}(c,z) = (\mathcal{K}^{c;\ell}\mathbf{R}^+ f)(z)$, since $\bar{\gamma} \leq \gamma_2 + \beta$, one has that $F^{(\ell)}(c,\cdot)$ is a modeled distribution with its $\|\!|\cdot|\!\|_{\bar{\gamma},\bar{\eta};t}$ norm bounded by $t^\kappa$, and by Corollary 4.3 the map $c \mapsto F^{(0)}(c,\cdot)$ is smooth (in the usual sense) into $\mathcal{D}_P^{\bar{\gamma},\bar{\eta}}$ with its derivatives given precisely by the $F^{(\ell)}$.

It then follows from the multiplicativity of the action of the structure group that

$$\begin{aligned}
&\sum \frac{(\hat{b}(x))^\ell}{\ell!} F^{(\ell)}(\bar{b}(x),x) - \Gamma_{xy}\Big(\sum \frac{(\hat{b}(y))^\ell}{\ell!} F^{(\ell)}(\bar{b}(y),y)\Big) \\
&\quad = \Big(\sum \frac{(\hat{b}(x))^\ell}{\ell!} F^{(\ell)}(\bar{b}(x),x) - \sum \frac{(\Gamma_{xy}\hat{b}(y))^\ell}{\ell!} F^{(\ell)}(\bar{b}(y),x)\Big) \\
&\qquad + \sum \frac{(\Gamma_{xy}\hat{b}(y))^\ell}{\ell!} \big(F^{(\ell)}(\bar{b}(y),x) - \Gamma_{xy}F^{(\ell)}(\bar{b}(y),y)\big) =: A_1 + A_2.
\end{aligned}$$

The term $A_1$ can be bounded precisely as in [7, prop. 6.13], with the only minor difference that the smooth function $F^{(\ell)}(\cdot,x)$ that $b$ is substituted into takes values in $T$ instead of $\mathbf{R}$. One then gets

$$\|A_1\|_m \lesssim t^\kappa \sum_{m_1+m_2=m} \|x-y\|_{\mathfrak{s}}^{\gamma_1-m_1} |x|_P^{\eta_1-\gamma_1} |x|_P^{(\bar\eta-m_2)\wedge 0},$$

where in the above sum $m_2$ runs over homogeneities of $\mathcal{I}\mathcal{W} + \overline{T}$; in particular, its smallest value is $(\alpha+\beta)\wedge 0 \geq \bar\gamma - \gamma_1$. Therefore,

$$\begin{aligned}\|A_1\|_m &\lesssim t^\kappa \sum_{m_1+m_2=m} \|x-y\|_{\mathfrak{s}}^{\bar\gamma-m} \|x-y\|_{\mathfrak{s}}^{m_2+\gamma_1-\bar\gamma} |x|_P^{\eta_1-\gamma_1} |x|_P^{(\bar\eta-m_2)\wedge 0} \\ &\lesssim t^\kappa \sum_{m_1+m_2=m} \|x-y\|_{\mathfrak{s}}^{\bar\gamma-m} |x|_P^{\eta_1-\bar\gamma} |x|_P^{\bar\eta\wedge m_2} \lesssim \|x-y\|_{\mathfrak{s}}^{\bar\gamma-m} |x|_P^{\bar\eta-\bar\gamma},\end{aligned}$$

where in the last step we used $\eta_1 \geq \bar\eta \vee 0$ and $\alpha+\beta > \bar\eta$. On the other hand,

$$\begin{aligned}\|A_2\|_m &\lesssim t^\kappa \sum_{m_1+m_2=m} \|x-y\|_{\mathfrak{s}}^{-m_1} \|x-y\|_{\mathfrak{s}}^{\bar\gamma-m_2} |x|_P^{\bar\eta-\bar\gamma} \\ &\lesssim t^\kappa \|x-y\|_{\mathfrak{s}}^{\bar\gamma-m} |x|_P^{\bar\eta-\bar\gamma}\end{aligned}$$

as required.

For a fixed model, bounding $|||\mathfrak{I}(b, \mathbf{R}^+ f);\ \mathfrak{I}(b, \mathbf{R}^+ \tilde f)|||_{\bar\gamma,\bar\eta;t}$ is immediate from the above thanks to the linearity of $\mathfrak{I}$ in the second argument. To bound

$$|||\mathfrak{I}(b, \mathbf{R}^+ f); \mathfrak{I}(\tilde b, \mathbf{R}^+ f)|||_{\bar\gamma,\bar\eta;t},$$

one can write, as in the proof of [7, thm. 4.16], with $b' = b - \tilde b$,

$$\begin{aligned}&\big(\mathfrak{I}(b, \mathbf{R}^+ f) - \mathfrak{I}(\tilde b, \mathbf{R}^+ f)\big)(x) = \\ &\qquad \sum_{\ell,i} \int_0^1 (b')_i(x) \frac{(\hat{\tilde b}(x) + \theta \hat b'(x))^\ell}{\ell!} (\mathcal{K}^{\bar{\tilde b}(x)+\theta \bar b'(x);\ell+e_i} \mathbf{R}^+ f)(x)\, d\theta,\end{aligned} \tag{4.5}$$

where the sum over $i$ runs over $1,\dots,d_1$, and $e_i$ is the $i^{\text{th}}$ unit vector in $\mathbb{R}^{d_1}$. Now one can repeat the preceding calculation, gaining a factor $|||b'|||_{\gamma_1,\eta_1;t}$ at each step.

Finally, to bound $|||\mathfrak{I}(b, \mathbf{R}^+ f);\ \mathfrak{I}(\tilde b, \mathbf{R}^+ \tilde f)|||_{\bar\gamma,\bar\eta;t}$ for two different models, one can employ the trick in [9, prop. 3.11]. $\square$

*Remark* 4.6. The same argument actually shows that if $c \mapsto F(c,\cdot)$ is a smooth function from $\mathfrak{K}$ to $\mathcal{D}_P^{\bar\gamma,\bar\eta}$ and $b = \bar b \mathbf{1} + \hat b$ is as in the statement, then the function $G$ given by

$$G(z) = \sum_{|\ell|\leq N} \frac{(\hat b(z))^\ell}{\ell!} F^{(\ell)}(\bar b(z), z)$$

belongs to $\mathcal{D}_P^{\bar\gamma,\bar\eta}$. This statement then has both the first part of Theorem 4.4 and [7, thm. 4.16] as corollaries.

To formulate the abstract counterpart of (2.5b), it remains to lift the operators $\hat{I}_k^{(\ell)}$. Using the notation

$$(K^\zeta u_0)(z) = \int K_t^\zeta(x-x')u_0(x')dx'$$

and identifying this function with its lift via its Taylor expansion, we define, similarly to $\mathfrak{I}$,

$$\hat{\mathfrak{I}}_k^{(m)}(b,u_0)(z) := \sum_{|\ell|\le N'} \frac{(\hat{b}(z))^\ell}{\ell!}\big(K^{\bar{b}(z);k+\ell}D_m u_0\big)(z). \tag{4.6}$$

Let us fix a noninteger exponent $1 > \eta_0 > \bar{\eta}$. This time the constant-coefficient result we rely on is the following variant of [7, lem. 7.5]:

LEMMA 4.7. *Assume $\beta = \mathfrak{s}_1$. Let $u_0 \in \mathcal{C}^{\eta_0}$ and $\zeta \in \mathcal{C}^{-N}$. Then $K^\zeta u_0 \in \mathcal{D}_P^{\gamma,\eta_0}$ for any $\gamma \ge \eta_0 \vee 0$, and*

$$\|K^\zeta u_0\|_{\gamma,\eta_0} \lesssim |\zeta|_{\mathcal{C}^{-N}}|u_0|_{\mathcal{C}^{\eta_0}}.$$

The behavior of $\hat{\mathfrak{I}}$ is then given by the following lemma.

LEMMA 4.8. *Assume $\beta = \mathfrak{s}_1$. Let $V$, $N'$, $b$, and $\tilde{b}$ be as in Theorem* 4.4, *and let $u_0, \tilde{u}_0 \in \mathcal{C}^{\eta_0}$. Then $\hat{\mathfrak{I}}(b,u_0) \in \mathcal{D}_P^{\bar{\gamma},\bar{\eta}}$ and with $\bar{\kappa} = (\eta_0 - \bar{\eta})/\mathfrak{s}_1$, one has the bounds, for any $t \in (0,1]$,*

$$\begin{aligned}
|\!|\!|\hat{\mathfrak{I}}(b,u_0)|\!|\!|_{\bar{\gamma},\bar{\eta};t} &\lesssim |\!|\!|b|\!|\!|_{\gamma_1,\eta_1;t}|u_0|_{\mathcal{C}^{\eta_0}},\\
|\!|\!|\hat{\mathfrak{I}}(b,u_0)\,;\hat{\mathfrak{I}}(\tilde{b},\tilde{u}_0)|\!|\!|_{\bar{\gamma},\bar{\eta};t} &\lesssim t^{\bar{\kappa}}\big(|\!|\!|b\,;\tilde{b}|\!|\!|_{\gamma_1,\eta_1;t} + \|(\Pi,\Gamma)-(\tilde{\Pi},\tilde{\Gamma})\|_{\bar{\gamma};2}\big)\\
&\quad + |u_0 - \tilde{u}_0|_{\mathcal{C}^{\eta_0}}.
\end{aligned}$$

*Moreover, the following identity holds*:

$$\mathcal{R}\hat{\mathfrak{I}}(b,u_0) = \hat{I}(\mathcal{R}b,u_0).$$

PROOF. The proof goes precisely as that of Theorem 4.4, with the only slight difference that the constant-coefficient estimate seemingly does not help in obtaining a positive power of $t$. Note, however, that whenever $\eta_0 > 0$, for any $c \in \mathfrak{K}$ and nonzero multi-index $\ell$, $\langle K^{c;\ell}u_0, \mathbf{1}\rangle$ vanishes at the initial time. In particular, whenever $\eta_0 < 1$, all components of $K^{c;\ell}u_0$ lower than $\eta_0$ vanish at the initial time, and hence (see [7, lem. 6.5]) one gets the estimate

$$\|K^{c;\ell}u_0\|_{\gamma,\bar{\eta}} \lesssim t^{\bar{\kappa}}|u_0|_{\mathcal{C}^{\eta_0}}.$$

It remains to notice that in the calculation for $|\!|\!|\hat{\mathfrak{I}}(b,u_0);\hat{\mathfrak{I}}(\tilde{b},u_0)|\!|\!|_{\bar{\gamma},\bar{\eta};t}$ analogous to (4.5), we only encounter instances of $K$ with nonzero derivatives with respect to the parameter $c$; hence the claimed factor $t^{\bar{\kappa}}$ in the lemma is indeed obtained. □

## 5 A Concrete Example

At this point, we have a completely automatic solution theory: given a quasilinear equation like (2.1), its solution is *defined* as $\mathcal{R}U$, where $U$ is obtained from the system of abstract equations

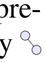
$$
\begin{aligned}
U &= \mathfrak{I}(a(U), \hat{\mathcal{F}}) + \hat{\mathfrak{I}}(a(U), u_0), \\
\hat{\mathcal{F}} &= (1 - V_3 a'(U)) F(U, \Xi) + 2V_1 a(U) a'(U) \mathscr{D}U \\
&\qquad + V_2 a(U)(a'(U))^2 (\mathscr{D}U)^2 + V_3 a(U) a''(U)(\mathscr{D}U)^2, \\
V_1 &= \mathfrak{I}_1'(a(U), \hat{\mathcal{F}}) + \hat{\mathfrak{I}}_1'(a(U), u_0), \\
V_2 &= \mathfrak{I}_2(a(U), \hat{\mathcal{F}}) + \hat{\mathfrak{I}}_2(a(U), u_0), \\
V_3 &= \mathfrak{I}_1(a(U), \hat{\mathcal{F}}) + \hat{\mathfrak{I}}_1(a(U), u_0).
\end{aligned}
\tag{5.1}
$$

If $F$ was a subcritical nonlinearity to begin with and $\alpha > -2$, then the above system is again subcritical, so one can use the construction of Section 3 to build the corresponding regularity structure. Provided it satisfies Assumption 3.8, one can use [4] in the form of Theorem 3.9 to obtain the corresponding BPHZ model. The local well-posedness of (5.1) is then a standard consequence of the results of Section 4 above just as in [7, sec. 6].

However, as mentioned in the introduction, at this point it is not automatic to see what counterterms appear (or whether they are even local in the solution) in the equation solved by $\mathcal{R}^\varepsilon U^\varepsilon$, where $U^\varepsilon$ is obtained from solving (5.1) with a renormalized smooth model. Below we carry out the computation of these terms in the setting of the example (1.1). An interesting outcome of these calculations is that if we consider the BPHZ renormalization of our model, then it may happen in general that nonlocal counterterms appear. However, as we will see, it is possible to choose the renormalization procedure in such a way that these nonlocal terms cancel out, thus leading to the stated result.

PROOF OF THEOREM 1.1. Our abstract equation reads as (5.1), with the nonlinearity $F(U, \Xi) = F_0(U)(\mathscr{D}U)^2 + F_1(U)\Xi$. The regularity structure is built as discussed above, where we declare the homogeneity of $\Xi$ to be $-3/2 + \kappa$ for some $\kappa \in (0, (\nu \wedge \bar{\nu})/2)$. As for the models, we take a slight modification of the associated BPHZ models $(\hat{\Pi}^\varepsilon, \hat{\Gamma}^\varepsilon)$ obtained from Theorem 3.9.

Recall first from Remark 3.10 that the BPHZ renormalization procedure is parametrized by functions $C_\tau^\varepsilon$ given by (3.8) for $\tau \in \mathcal{W}_- := \{\tau \in \mathcal{W} \setminus \{\mathbf{1}\} : |\tau| \leq 0\}$. As a consequence of the fact that we choose $|\Xi| > -\frac{3}{2}$, one can verify that all $\tau \in \mathcal{W}_-$ satisfy $[\![\tau]\!] \leq 3$. Since we furthermore assumed that the driving noise $\xi$ is centered Gaussian, the functions $C_\tau^\varepsilon$ vanish identically for all $\tau$ with $[\![\tau]\!]$ odd, so that only symbols with $[\![\tau]\!] = 2$ contribute to the renormalization.

Using the graphical notation from [8, 9] (circles represent $\Xi$, plain lines represent $\mathcal{I}$, and bold red lines represent $\mathcal{I}'$), the only two such symbols are given by 

and . The corresponding renormalization functions are given by

$$
\begin{aligned}
C^{\varepsilon}(c) &= \int K^{(c)}(z)\mathscr{C}^{\varepsilon}(z)\,dz,\\
C^{\varepsilon}(c,\bar c) &= \int \partial_x K^{(c)}(z)\partial_x K^{(\bar c)}(\bar z)\mathscr{C}^{\varepsilon}(\bar z - z)\,dz\,d\bar z,
\end{aligned}
\tag{5.2}
$$

where $\mathscr{C}^{\varepsilon} = \mathscr{C} * \rho^{\varepsilon} * \rho^{\varepsilon}$. In this particular case, this allows us as in [3] to define linear maps $M^{\varepsilon}\colon T \to T$ such that on $T_{\leq 0}$ the BPHZ renormalized model $(\hat\Pi^{\varepsilon}, \hat\Gamma^{\varepsilon})$ satisfies the identity

$$
\hat\Pi^{\varepsilon}_z \tau = \Pi^{\varepsilon}_z M^{\varepsilon}\tau,
\tag{5.3}
$$

where $\Pi^{\varepsilon}$ is the canonical lift of $\xi^{\varepsilon}$. (The fact that (5.3) holds is no longer the case when $\kappa \leq 0$!) One has, for example,

$$
\begin{aligned}
M^{\varepsilon}(\zeta \otimes ) &= \zeta \otimes  - \zeta(C^{\varepsilon})\mathbf{1},\\
M^{\varepsilon}(\zeta\otimes\eta\otimes\nu\otimes ) &= \zeta\otimes\eta\otimes\nu\otimes  - (\zeta\otimes\eta)(C^{\varepsilon})\,\nu\otimes .
\end{aligned}
\tag{5.4}
$$

At this point we note that if $K^{(c)}$ were exactly equal to the heat kernel instead of a compactly supported truncation, then one would have the exact identity

$$
C^{\varepsilon}(c) = c\,C^{\varepsilon}(c,c).
\tag{5.5}
$$

It turns out that this identity is crucial in order to obtain the cancellations necessary to obtain local counterterms. We therefore define a model $(\tilde\Pi^{\varepsilon}, \tilde\Gamma^{\varepsilon})$ just like the BPHZ model, but with $C^{\varepsilon}$ *defined* by (5.5) instead of (5.2). Since the difference between these two different definitions of $C^{\varepsilon}$ converges to a finite smooth function as $\varepsilon \to 0$, the convergence of the BPHZ model as $\varepsilon \to 0$ also implies the convergence of the model $(\tilde\Pi^{\varepsilon}, \tilde\Gamma^{\varepsilon})$.

Note also that (modulo changing the order of some factors: recall that  $\neq$  in our setting, but this distinction is essentially irrelevant since we always consider models such that for example $\Pi_x(\zeta\otimes\eta\otimes\nu\otimes ) = \Pi_x(\nu\otimes\zeta\otimes\eta\otimes )$, so that we can identify such elements for all practical purposes), one has

$$
\mathcal{W}_- = \{\ , X_2, \dots\}.
\tag{5.6}
$$

Inspecting (5.4), as well as the analogous expressions for all other symbols of negative degree, we conclude that one has

$$
(\tilde\Pi^{\varepsilon}_z\tau)(z) = (\Pi^{\varepsilon}_z\tau)(z) + \langle \mathbf{1}, (M^{\varepsilon} - \mathrm{id})\tau\rangle
\tag{5.7}
$$

for all $\tau \in T$, where $\langle \mathbf{1}, \sigma\rangle$ denotes the coefficient of $\mathbf{1}$ in $\sigma$. Furthermore, the second term in this expression is nonvanishing only if $\tau$ contains a summand in $T$ and/or in $T$. This is because of (5.3), combined with the fact that $M^{\varepsilon}$ only generates terms of strictly positive degrees for the remaining symbols in $\mathcal{W}$.

We now have everything in place to derive the form of the renormalized equation. Given the $(\tilde\Pi^{\varepsilon}, \tilde\Gamma^{\varepsilon})$ for some fixed $\varepsilon > 0$, one obtains a local solution of the system (5.1) in

$$
\begin{aligned}
&\mathcal{D}^{3/2+2\kappa,2\kappa}(W_0) \oplus \mathcal{D}^{\kappa,-2+3\kappa} \oplus \mathcal{D}^{1/2+2\kappa,-1+2\kappa}(W_1) \\
&\quad \oplus \mathcal{D}^{1+2\kappa,2\kappa}(W_0) \oplus \mathcal{D}^{3/2+2\kappa,2\kappa}(W_0),
\end{aligned}
\tag{5.8}
$$

where $W_0$ is the sector generated by the Taylor polynomials and elements of the form $\mathcal{I}^{\zeta}\tau$, and $W_1 = \mathscr{D}W_0$. As a consequence of (5.7), we conclude as for example in [7, sec. 9.3] that for $\varepsilon > 0$ the pair $(\mathcal{R}^{\varepsilon}U^{\varepsilon}, \mathcal{R}^{\varepsilon}\widehat{\mathcal{F}}^{\varepsilon})$ solves an equation just like (2.5), but with an additional term $\langle \mathbf{1}, (M^{\varepsilon} - \mathrm{id})\widehat{\mathcal{F}}^{\varepsilon}\rangle$ appearing on the right-hand side of (2.5b). Hence $\mathcal{R}^{\varepsilon}U^{\varepsilon}$ solves an equation just like (2.1), but with an additional term

$$
\mathcal{E} := \frac{\langle \mathbf{1}, (M^{\varepsilon} - \mathrm{id})\widehat{\mathcal{F}}^{\varepsilon}\rangle}{1 - a'(u)\langle \mathbf{1}, V_3^{\varepsilon}\rangle}
\tag{5.9}
$$

appearing on the right-hand side.

It now remains to show that if $(U, \widehat{\mathcal{F}})$ solves (5.1), then (5.9) coincides with a local functional of $u = \mathcal{R}U = \langle \mathbf{1}, U\rangle$. Furthermore, write $v_i = \mathcal{R}V_i = \langle \mathbf{1}, V_i\rangle$, as well as $q = 1 - v_3 a'(u)$, where the $V_i$ are as in (5.1). Note that $q$ is the denominator in (5.9), and that this is *not* a local functional of $u$, so that we should aim for a factor $q$ to appear in the numerator as well. To ease notation, we henceforth also omit the lower indices in $\delta_{a(u)}$ and $\delta'_{a(u)}$. Since all symbols appearing in the expansion of the solution are of the form $\zeta \otimes \tau$, where $\zeta$ is a tensor product of either $\delta_{a(u)}$ or $\delta'_{a(u)}$, this will hopefully not cause any confusion.

To calculate the numerator in (5.9), it follows from the above discussion that we only need to know the components of $\widehat{\mathcal{F}}$ in $T_{\circ\!-\!\circ}$ and in $T_{\circ\vee\circ}$. For this, note first that one has

$$
\widehat{\mathcal{F}} = qF_1(u)\,\circ + (\cdots),
\tag{5.10}
$$

where all terms included in $(\cdots)$ are of strictly higher degree than that of $\Xi$. Combining (5.1) with the definitions of $\mathfrak{I}$ and $\mathfrak{I}_1$, we then see that

$$
U = u\,\mathbf{1} + u_{\overset{\circ}{\bullet}} \otimes \overset{\circ}{\bullet} + \widetilde{U},
\tag{5.11}
$$

where $\widetilde{U}$ takes values only in spaces $T_\tau$ with $\tau \neq \overset{\circ}{\bullet}$ of the form $\tau = \prod_i \mathcal{I}^{\zeta_i}(\sigma_i)$. Furthermore, by (5.10) and the definition of $V_3$, the distribution $u_{\overset{\circ}{\bullet}}$ is given by

$$
u_{\overset{\circ}{\bullet}} = qF_1(u)\delta + a'(u)v_3 u_{\overset{\circ}{\bullet}} \quad \Rightarrow \quad u_{\overset{\circ}{\bullet}} = F_1(u)\delta,
\tag{5.12}
$$

so that in particular

$$
a(U) = a(u)\,\mathbf{1} + (a'F_1)(u)\delta \otimes \overset{\circ}{\bullet} + (\cdots).
\tag{5.13}
$$

Combining this with (5.10) and the expressions for $V_i$, we conclude similarly that

$$
\begin{aligned}
&V_1 = qF_1(u)\delta' \otimes \overset{\circ}{\bullet} + v_1\,\mathbf{1} + (\cdots), \quad V_2 = v_2\,\mathbf{1} + (\cdots), \\
&V_3 = v_3\mathbf{1} + (qF_1(u)\delta' + v_2(a'F_1)(u)\delta) \otimes \overset{\circ}{\bullet} + (\cdots),
\end{aligned}
$$

where the terms denoted by $(\cdots)$ are of higher degree. Combining all of this with the expression for $\widehat{\mathcal{F}}$ in (5.1), we finally obtain the next order in the development of $\widehat{\mathcal{F}}$, namely,

$$\begin{aligned}\widehat{\mathcal{F}} = {} & qF_1(u)\circ + \big(q(F_1'F_1)(u) - v_3(a''F_1^2)(u) - v_2((a')^2F_1^2)(u)\big)\,\delta\otimes{\circ\!\!\searrow\!\circ} \\ & - q(a'F_1^2)(u)\delta'\otimes{\circ\!\!\searrow\!\circ} + 2q(aa'F_1^2)(u)\,\delta'\otimes\delta\otimes{{}^{\circ}\!\vee\!{}^{\circ}} \\ & + \big(q(F_1^2F_0)(u) + v_2(a(a')^2F_1^2)(u) + v_3(aa''F_1^2)(u)\big)\delta\otimes\delta\otimes{{}^{\circ}\!\vee\!{}^{\circ}} + (\cdots).\end{aligned}$$

Combining this with the definition of $M^{\varepsilon}$, we conclude that the counterterm (5.9) is given by

$$\begin{aligned}\mathcal{E} = {} & -\frac{1}{q}\big(q\big(F_1'F_1\big)(u) - v_3\big(a''F_1^2\big)(u) - v_2((a')^2F_1^2)(u)\big)C^{\varepsilon}_{\circ\!\searrow\!\circ}(a(u)) \\ & + \big(a'F_1^2\big)(u)(\partial C^{\varepsilon}_{\circ\!\searrow\!\circ})(a(u)) - 2\big(aa'F_1^2\big)(u)(\partial_1 C^{\varepsilon}_{{}^{\circ}\!\vee\!{}^{\circ}})(a(u),a(u)) \\ & - \frac{1}{q}\big(q\big(F_1^2F_0\big)(u) + v_2(a(a')^2F_1^2)(u) + v_3\big(aa''F_1^2\big)(u)\big)C^{\varepsilon}_{{}^{\circ}\!\vee\!{}^{\circ}}(a(u),a(u)).\end{aligned}$$

At this point we note that by (5.5) and the fact that $C^{\varepsilon}_{{}^{\circ}\!\vee\!{}^{\circ}}$ is symmetric in its two arguments, one has the identity

$$(\partial C^{\varepsilon}_{\circ\!\searrow\!\circ})(c) = C^{\varepsilon}_{{}^{\circ}\!\vee\!{}^{\circ}}(c,c) + 2c(\partial_1 C^{\varepsilon}_{{}^{\circ}\!\vee\!{}^{\circ}})(c,c). \tag{5.14}$$

Inserting this into the above equation and noting that the terms proportional to $v_2$ and $v_3$ cancel out exactly thanks again to (5.5), we conclude that

$$\mathcal{E} = -C^{\varepsilon}_{\circ\!\searrow\!\circ}(a(u))\big(\big(aF_1'F_1\big)(u) + \big(F_1^2F_0\big)(u) - \big(a'F_1^2\big)(u)\big)/a(u), \tag{5.15}$$

which is precisely as in (1.2). □

*Remark* 5.1. The expression (5.2) also gives some information about the behavior of $C^{\varepsilon}_{\circ\!\searrow\!\circ}$ in the case where $\mathscr{C}$ is self-similar on small scales, i.e., $\mathscr{C}(\lambda^2 t, \lambda x) = \lambda^{-3+\nu}\mathscr{C}(t,x)$ for all $\lambda \in (0,1]$ and $|t|^{1/2} + |x| \le r$ for some $r > 0$. Indeed, one can then write

$$\begin{aligned}C^{\varepsilon}_{\circ\!\searrow\!\circ}(c) &= \int K^{(c)}(z)\mathscr{C}^{\varepsilon}(z)\,dz \approx \varepsilon^{-3+\nu}\int K^{(c)}(z)\mathscr{C}^1(\varepsilon z)\,dz \\ &= \varepsilon^{\nu}\int K^{(c)}(\varepsilon^{-1}z)\mathscr{C}^1(z)\,dz \approx \varepsilon^{-1+\nu}\int K^{(c)}(z)\mathscr{C}^1(z)\,dz,\end{aligned}$$

where $\approx$ means that the difference of the two sides converge as $\varepsilon \to 0$ to a smooth function of $c$. Hence, modulo changing again the renormalization constants by a finite quantity, one can use in this case a counterterm of the form $\varepsilon^{\nu-1}A(u)$ for some explicit function $A$ of the solution $u$.

**Acknowledgment.** MH gratefully acknowledges support by the Leverhulme Trust and by an ERC consolidator grant, project 615897.

## Bibliography


[1] Bailleul, I.; Debussche, A.; Hofmanová, M. Quasilinear generalized parabolic Anderson model equation. Preprint, 2016. arXiv:1610.06726 [math.AP]

[2] Bruned, Y.; Chandra, A.; Chevyrev, I.; Hairer, M. Renormalizing SPDEs in regularity structures. Preprint, 2017. arXiv:1711.10239 [math.AP]

[3] Bruned, Y.; Hairer, M.; Zambotti, L. Algebraic renormalisation of regularity structures. Preprint, 2016. arXiv:1610.08468 [math.RA]

[4] Chandra, A.; Hairer, M. An analytic BPHZ theorem for regularity structures. Preprint, 2016. arXiv:1612.08138 [math.PR]

[5] Furlan, M.; Gubinelli, M. Paracontrolled quasilinear SPDEs. Preprint, 2016. arXiv:1610.07886 [math.PR]

[6] Gerencsér, M.; Hairer, M. Singular SPDEs in domains with boundaries. Preprint, 2017. arXiv:1702.06522 [math.PR]

[7] Hairer, M. A theory of regularity structures. *Invent. Math.* **198** (2014), no. 2, 269–504. doi:10.1007/s00222-014-0505-4

[8] Hairer, M. The motion of a random string. Preprint, 2016. arXiv:1605.02192 [math.PR]

[9] Hairer, M.; Pardoux, É. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan* **67** (2015), no. 4, 1551–1604. doi:10.2969/jmsj/06741551

[10] Otto, F.; Sauer, J.; Smith, S.; Weber, H. Parabolic equations with low-regularity coefficients and rough forcing. In preparation.

[11] Otto, F.; Weber, H. Quasilinear SPDEs via rough paths. Preprint, 2016. arXiv:1605.09744 [math.AP]



MÁTÉ GERENCSÉR
IST Austria
Am Campus 1
3400 Klosterneuburg
AUSTRIA
E-mail: mate.gerencser@ist.ac.at

MARTIN HAIRER
Imperial College London
SW7 2AZ London
UNITED KINGDOM
E-mail: m.hairer@imperial.ac.uk


Received December 2017.